

CHAPTER 13 CHARTING IN EXCEL

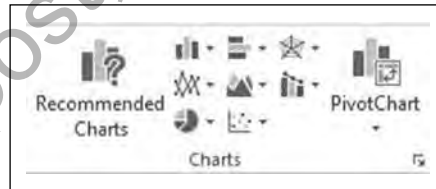
Now we will use Excel to create the charts that we saw in Chapter 12. Download the charting.xlsx file from the website for this book and open it. It contains seven worksheets, which have labels on their tabs. The first one is Degrees conferred. This sheet, like all of the others, has data that have been edited and formatted so it's easier to create the charts. Often, the data we want to use for our charts are stored in columns or rows that aren't neighbors. That makes it challenging to highlight just what we need. The data sources are noted at the bottom of each sheet and include URLs for download.

PIE CHART

We'll start by creating a pie chart to show the proportion of degrees conferred in 2009–2010, using data from the National Center for Education Statistics. The first two rows hold the data we need to build the chart. It has the degree category in the first row and the numbers for each category in the second.

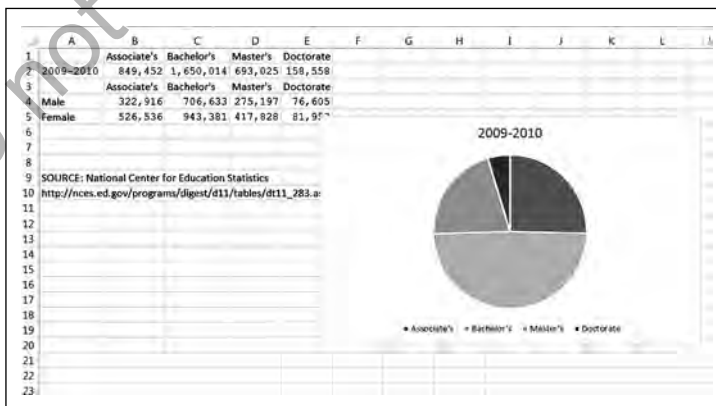
Highlight cells A1 through E2. Then select the Insert tab and click on the drop-down arrow for the pie chart button.

Select the 2-D Pie option and Excel creates the chart just like that.



Source: Microsoft Excel for Windows 2013.

Note: Excel chart type selector.

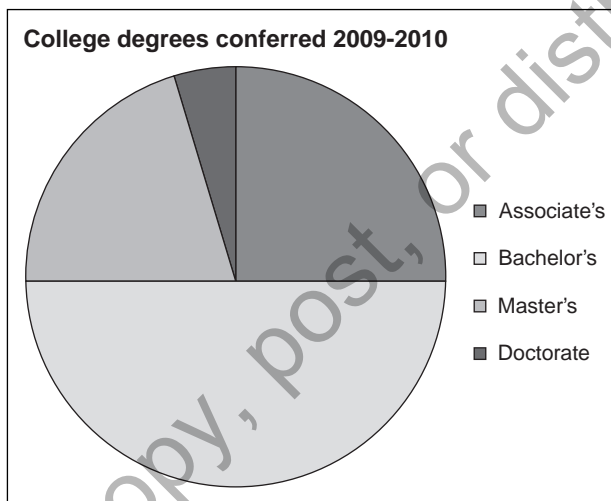


Source: National Center for Education Statistics, Department of Education.

Note: Creating a pie chart.

The chart is not very descriptive as it stands, so we are going to add some elements that will help viewers make sense of it. Double-click on the chart title at the top; now we can edit it. Let's change the text to, "College degrees conferred 2009–2010". The legend, which contains the color key for each degree, is at the bottom of the chart. Let's move that to the right of our pie. Click on the legend to resize and move it.

Finally, we'll add a source line below the pie so we can cite the NCES and provide a URL to the data. Click on the Format tab and then look for the Insert Shapes tool on the left. Select the Text Box tool and then use your cursor to enter the source information on the first line and the URL on the second. Resize the pie chart area if there is not enough room for the source line. Our finished chart should look like this:



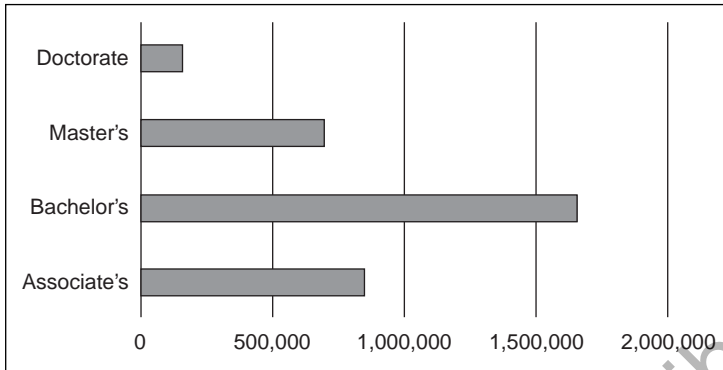
Source: National Center for Education Statistics, Department of Education.

Note: Pie chart with title, legend and source line.

Examining this chart, we can see that nearly half of all degrees awarded were bachelor's degrees. Associate's degrees make up about one quarter; master's and doctorate degrees, combined, make up another quarter. We've done some work, so let's create a working copy of this file by picking File | Save As.

HORIZONTAL BAR CHARTS

We can also use horizontal bar charts if we want to compare our degree categories. Once again, highlight cells A1 through E2. Pick Insert Bar Chart, then select 2-D Bar, Clustered Bar. Excel creates this chart. We could edit this if we wanted by modifying the chart

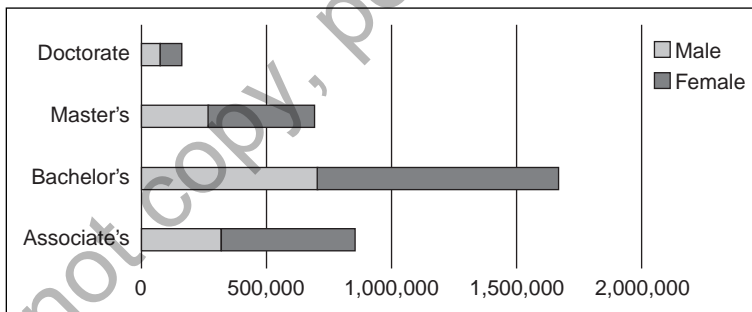


Source: National Center for Education Statistics, Department of Education.

Note: Horizontal bar chart.

title and adding a source line. Excel has automatically chosen to use degree numbers in increments of 500,000 for the horizontal axis labels.

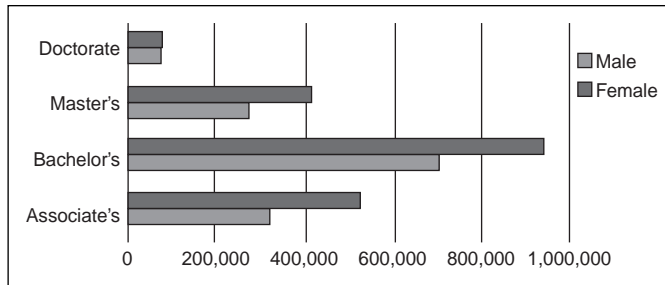
Right below our table we have another that breaks down the degrees earned by sex. Highlight cells A4 to E6. Insert a Bar Chart and pick Stacked. That will create the same chart with four bars, this time showing the shares earned by women and by men.



Source: National Center for Education Statistics, Department of Education.

Note: Stacked horizontal bar chart shows proportions.

As we saw in Chapter 12, the clustered bar chart allowed us to see the differences more clearly. So let's build that. Instead of creating a new chart from scratch, we can right-click on the chart area and then pick Change Chart Type. . . . Pick Clustered Bar from the box that appears and Excel creates this chart. It's much easier to see the differences between the sexes with this one.



Source: National Center for Education Statistics, Department of Education.

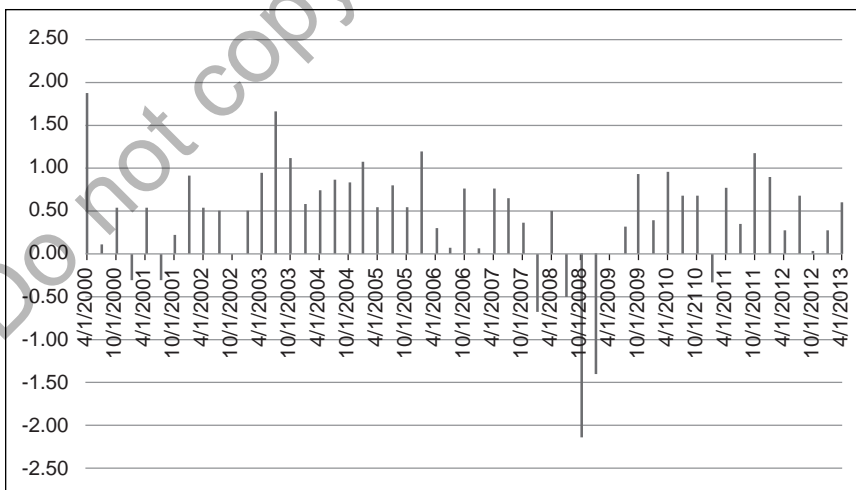
Note: Clustered horizontal bar chart compares categories.

COLUMN AND LINE CHARTS

Now that we've charted categorical data from one point in time, we are going to learn how to chart time-series data. Remember that we have two choices when charting time-series data: (1) column charts for data that were recorded or reported at discrete points in time, and (2) line charts for data that measure a continuous phenomenon, such as the weather.

We'll use the GDP data on the GDP sheet to create our first column chart. The GDP data are reported quarterly, as reflected by the dates stored in column A. Column B tells us the percentage change from the previous quarter.

Highlight A1 through B54 to grab the data that we need to create our chart. Insert a 2-D Clustered Column chart. Excel creates a barebones chart that lacks a lot of descriptive information. Also, we see that our horizontal axis labels have been stuffed into the middle of the chart, making them difficult to read.



Source: Department of Commerce, retrieved from FRED.

Note: Vertical column chart.

We're going to clean this chart up and discover some tricks that we can use later. First, let's fix the position of the horizontal axis labels. Select the labels and then go to the Format Axis Box that appears. Pick the column chart button and open the Labels item. Use the drop-down to pick Low as the label position.

Now we will label both of our axes, so chart viewers can understand what each means. Pick Add Chart Element from the Design tab and then Axis Titles | Primary Vertical. A text box appears. Enter "Percent change". Likewise, create a horizontal axis label that says "Quarter". Retitle the chart to "Quarterly change, U.S. real GDP 2000–2013".

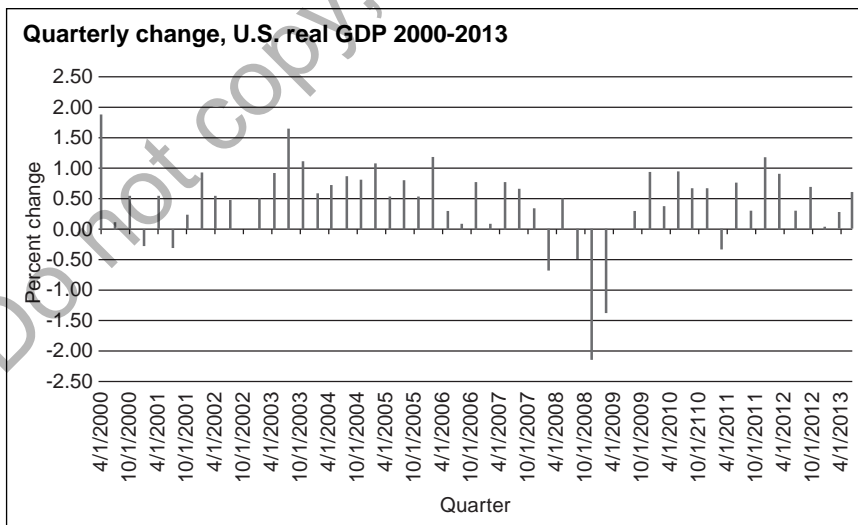
Last, we want to insert a source line. Right now we don't have enough room at the bottom of the chart. We can resize the chart plot area to free up some space. Click on the plot area and resizing handles appear. Drag them to make the plot area smaller.

Now it looks much better, and someone can understand what it means.



Source: Microsoft Excel for Windows 2013.

Note: Format axis settings.

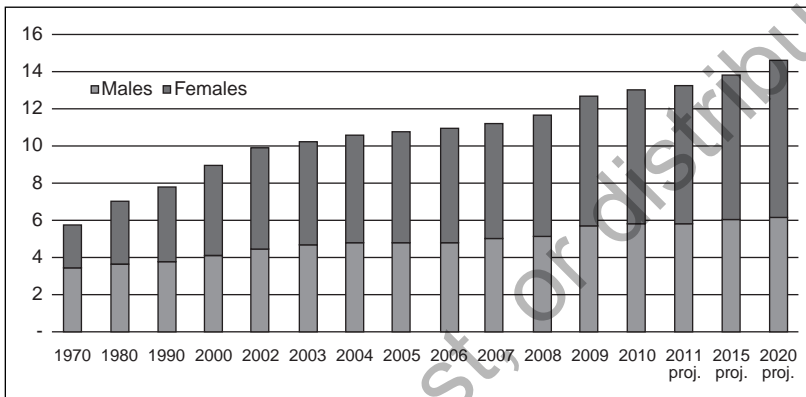


Source: Department of Commerce, retrieved from FRED.

Note: Category labels placed at bottom of axis.

Our next column charts will be those in which we'll display fall enrollment at U.S. degree-granting institutions of higher education over time. Go to the Enrollment tab to find these data from the NCES. Row 1 holds the categories for years. Rows 2 and 3 hold enrollment data that the author has reformatted so the vertical axis labels appear properly when charting. The numbers are in millions, rounded up to whole numbers. The original numbers appear in Rows 4 and 5 and are reported in thousands.

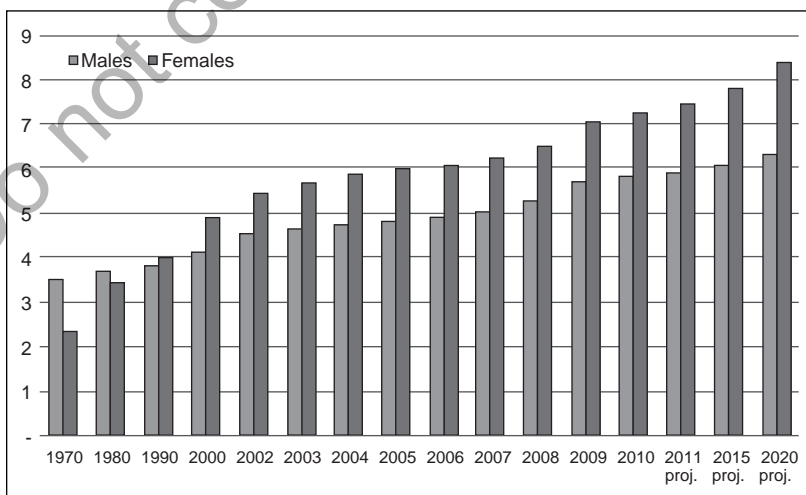
Highlight cells A1 through S3 and insert a Stacked Column chart to create this chart. Our chart shows increased enrollment since the 1970s, with women increasing their share.



Source: National Center for Education Statistics, Department of Education.

Note: Stacked vertical column chart.

Use Change Chart Type and pick Clustered Column. We see that more men than women had enrolled in 1970. Sometime in the 1980s women surpassed men. We can see that the current gap is expected to widen even more during the coming years.

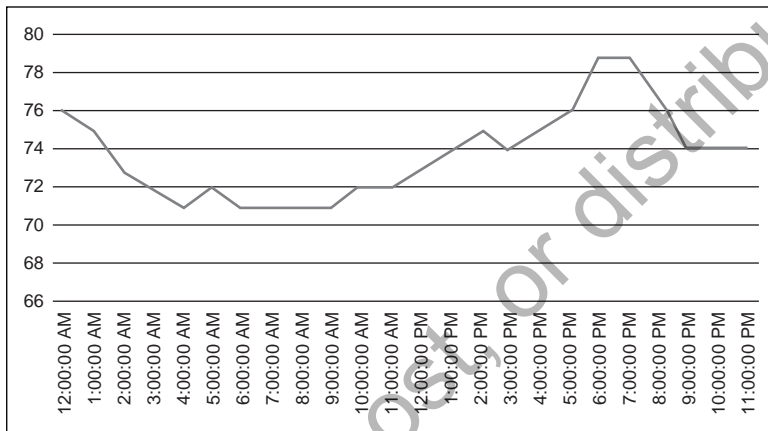


Source: National Center for Education Statistics, Department of Education.

Note: Clustered vertical column chart.

Let's turn to charting using continuous time series data. Go to the PBGH Weather sheet, which has hourly temperature and relative humidity data from September 1, 2013; the author downloaded these data from the National Weather Service. First, we will plot just one variable, the temperature. Then we will add the humidity readings.

Select cells A1 through B25 and insert a 2-D Line Chart. This chart appears. It looks like the temperatures bounced around a lot on that date. That's because the bottom value on the vertical axis is 66 degrees. We can set that to a lower value to render a more accurate picture of the temperature on that date.



Source: National Weather Service.

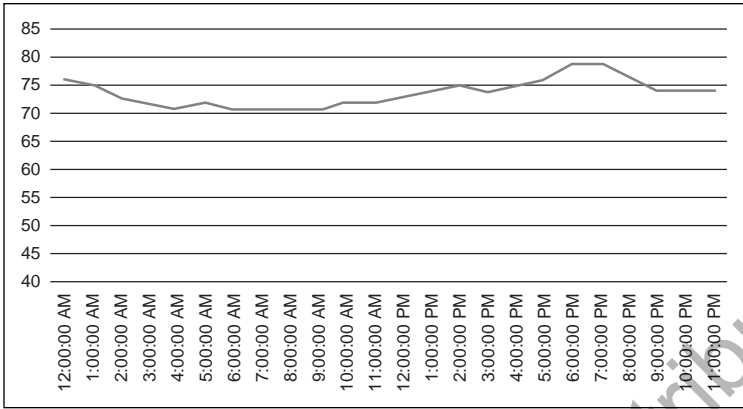
Note: Line chart. Note that the temperatures appear more extreme than they were because the vertical axis starts at 66 degrees.

Select the vertical axis labels and then, in the Format Axis box, set the Bounds Minimum to 40. The chart now looks much better.



Source: Microsoft Excel for Windows 2013.

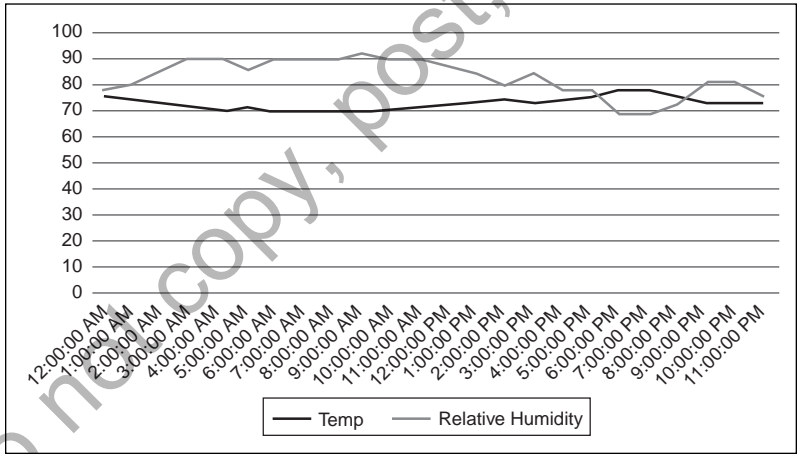
Note: Modify axis in vertical axis setting.



Source: National Weather Service.

Note: Temperature variation smoothed, thanks to vertical axis minimum lowered to 40 degrees.

Next we will create a chart that has the temperature and relative humidity. Select Cells A1 through C25 and insert a 2-D Line Chart. We get this chart, which shows both, plotted on the same 0- to 100-point scale. Of course, we could modify that scale if we wanted to.

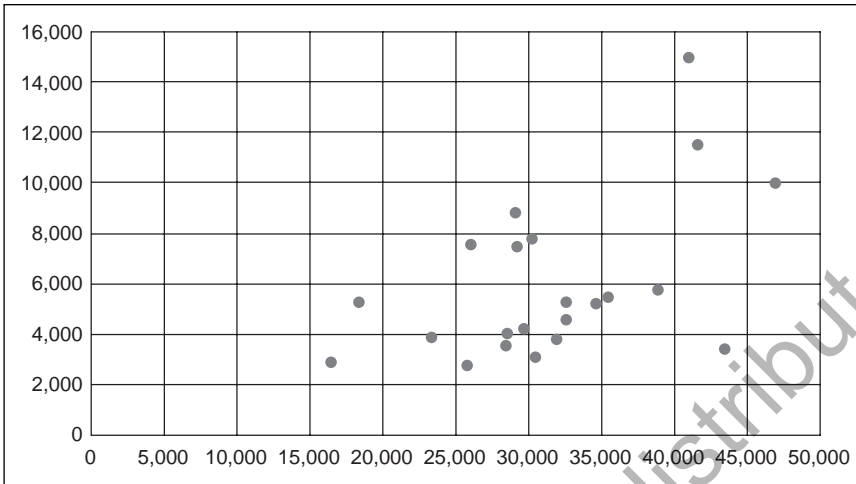


Source: National Weather Service.

Note: Line for relative humidity added to chart.

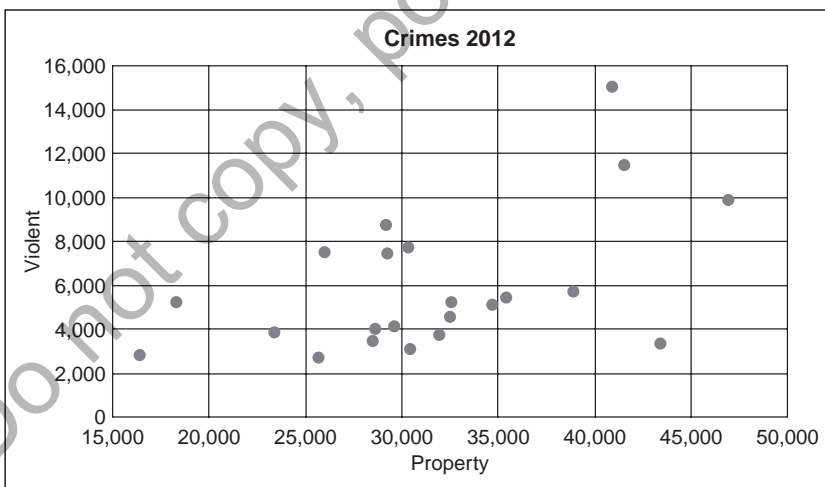
SCATTERPLOT

For our next chart we will again plot two variables. As we learned in the previous chapter, scatterplots can help us see whether there is a relationship between numeric variables. Go to the Crimes worksheet, which has data about property crimes and violent crimes in cities with populations 500,000 to 999,999. Highlight the cells containing the data—B1 through C23—and then pick Insert Scatter, using the first option. From the distribution of the points, it does look like there is a relationship between the number of property and violent crimes.



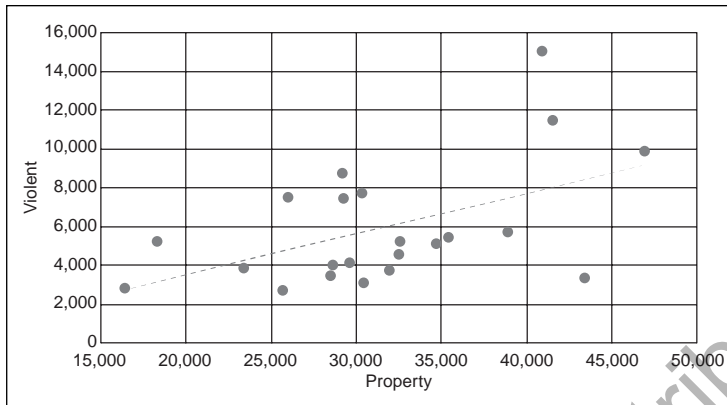
Source: Federal Bureau of Investigation.
 Note: Scatterplot of property and violent crimes.

Change the title to “Crimes 2012”. Label the horizontal axis “Property” and the vertical axis “Violent”. No city has fewer than 16,000 property crimes, so change the minimum bounds for that axis to 15000. Our chart looks better now.



Source: Federal Bureau of Investigation.
 Note: Scatterplot with new horizontal axis settings and labels.

Let’s add a finishing touch, a linear trend line. Excel uses a statistical function called linear regression to draw the line that’s a best fit for the data. Points that are very high above the line represent cities with more violent crimes than we would expect, based on their numbers of property crimes.

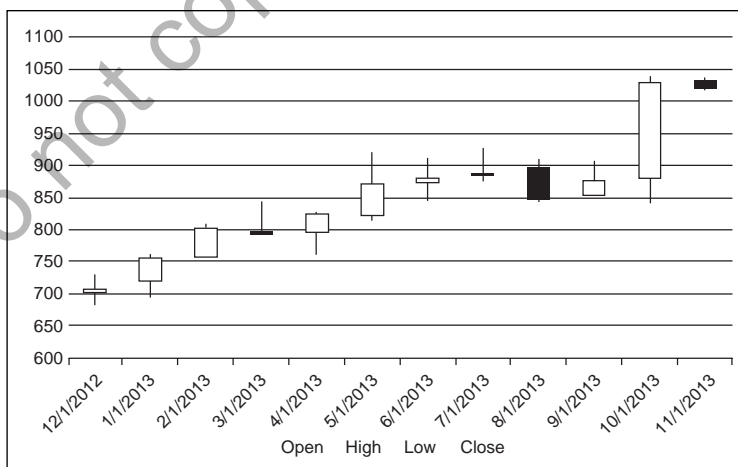


Source: Federal Bureau of Investigation.

Note: Scatterplot with linear trend line showing the central tendency.

STOCK CHART

Pick the Google stock sheet, which contains monthly performance data downloaded from Yahoo Finance. Note that when we create stock charts our data need to be ordered properly: date, opening price, high price, low price and closing price. Date represents the first day of trading for each month. Highlight cells A1 through E15 and Insert Stock, Surface or Radar Chart (the one that looks like a spiderweb). Then pick the Open-High-Low-Close option. Change the vertical axis label's minimum bounds to 600 to eliminate the empty space at the bottom. This also allows us to better distinguish how the stock performed in each month. We see that some of our inner bars, which display the closing and opening prices, are filled in. That's Excel's way of telling us that the stock closing price fell below its opening price in those months.



Source: Yahoo Finance.

Note: Stock chart showing opening, closing, high and low prices.

SPARKLINES

We'll wrap up this chapter by learning how to create sparklines. Pick the Median income tab to find data from the NCES. We have nine columns of salary amounts in different years and more than 50 categories. The procedure for creating sparklines is a little more involved than for the other charts we've created so far. Click on the Line button in the Sparklines group and then set the Data Range to B2:J67. Next tell Excel where to put the sparklines. The Location Range needs to be K2:K67. Click OK and we've placed our sparklines in the K column.

2	All persons, all education levels	\$36,900	\$39,190	\$37,990	\$38,140	\$37,970	\$37,600	\$39,600	\$38,560	\$37,950
3	Less than high school completion	23,480	23,640	23,660	22,320	23,780	22,330	21,850	21,660	22,860
4	High school completion\3\	30,680	32,650	32,130	32,290	31,420	31,310	31,370	30,690	29,390
5	Some college, no degree	34,310	37,680	36,340	35,020	35,660	35,390	34,780	34,940	31,990
6	Associate's degree	36,630	39,180	39,120	37,720	37,750	37,540	37,640	38,180	37,030
7	Bachelor's or higher degree	48,670	52,240	50,530	50,080	51,910	52,180	52,290	50,370	50,000
8	Bachelor's degree	45,700	52,130	46,990	46,320	48,630	48,060	47,100	46,420	44,970
9	Master's or higher degree	58,520	62,590	57,520	55,790	60,350	57,400	62,090	56,350	55,230

Source: National Center for Education Statistics, Department of Education.

Note: Sparklines.

That's it for our lesson on making Excel charts. In the next chapter, we'll wrap up this section on data visualization with a look at Web-based tools that allow us to create and share charts.

ON YOUR OWN

Download the Excel spreadsheet charting2.xlsx from the book website and use the data provided on each tab to create the best chart possible. Make a note of why you chose the particular chart type.